PTO/PCT Rec'd 24 AUG 2001
PCT/SG99/00014
09/914171

METHOD AND APPARATUS FOR INTERLACED/NON-INTERLACED FRAME DETERMINATION, REPEAT-FIELD IDENTIFICATION AND SCENE-CHANGE DETECTION

The present invention relates to methods and apparatus for the pre-processing of moving pictures before encoding. In particular, the present invention relates to methods and apparatus for determining whether a digital picture frame is an interlaced-scan picture or a non-interlaced-scan picture; identifying a repeated-field; and detecting a scene-change in a sequence of moving pictures.

Encoding methods such as the well known MPEG-1 and MPEG-2 standards have been popularly used for efficient transmission and storage of video. An MPEG encoder compresses an input video signal picture-by-picture to produce an output signal or bitstream compliant to the relevant MPEG standard. Pre-processing techniques can be applied to the input video signal before encoding, for example, to remove noise and re-format the signal (eg. 4:2:2 to 4:2:0 conversion, image size conversion, etc.).

The input video signal is typically in an interlaced format, for example the 525/60 or 625/50 (lines/frequency) format, with each video frame consisting of two fields (top field and bottom field). However, the source material of the video signal may be originally produced on film and converted to the video signal via a telecine process. This process converts a progressive source into an interlaced format and provides at the same time, if necessary, frame rate conversion for example using a 3:2 or 2:2 pulldown technique. In the case of 24 Hz film to 525/60 Hz video conversion, each progressive film picture is converted to two interlaced video fields and, in addition, there are 12 repeated fields according to the 3:2 pulldown patterns in every second of the converted video. Improvement in coding efficiency can be obtained if the video source from film is identified and the repeated (or redundant) fields are detected and removed before coding. Pre-processing techniques applied before encoding can also gain from the results of film picture detection.

The known methods of film mode detection can be widely classified into two categories: (1) film mode detection using film-frame pattern identification; and (2) film mode detection using automatic interlace/progressive frame detection.

The output of the type of method using film-frame pattern identification is a decision whether the input sequence is an interlaced video or a 3:2/2:2 pulldown film. The detection tries to identify the unique pattern of a 3:2 or 2:2 pulldown film. One of the most commonly used techniques is to detect the repeat field pattern in the 3:2 pulldown film (as described in US patents 5,317,398 and 5,398,071). The pixel to pixel field differences between alternate fields (fields with the same parity) are measured to identify whether the 3:2 repeat field pattern exists.

Another commonly used assumption is that the field differences between two interlaced fields is significantly greater than the field difference between two non-interlaced (or progressive) fields. One method is to group the successive fields that have the least field differences as a film frame (as described in US patent 5,565,998). Another method is to measure the consecutive field differences of incoming fields and monitor the pattern to decide if it is an interlaced video, 3:2 film or 2:2 film (as described in US patents 5,365,273 and 5,689,301). In the above methods, the unique pattern is monitored for a period (typically spanning 5 to 64 fields) before a decision is made.

With the method of film mode detection using automatic interlace/progressive frame detection, apart from deciding whether an incoming sequence is a film, this type of detection also determines if a frame is interlaced or progressive and identifies a repeated field. Due to the inclusion of the interlace/progressive detection for every frame, it does not have the slow response in interlace/progressive encoding as in the film-frame pattern identification methods described above. One of the methods used for the interlace/progressive detection, such as in US patent No. 5,452,011, is the intra-field and inter-field difference (IIFD) comparison. The IIFD method compares the inter-field and intra-field differences to detect whether two consecutive fields are interlaced. The assumption is that the inter-field difference will be

greater than the intra-field difference.

In most of the current video/film detection methods which have no automatic interlace/progressive detection, when there is a transition from interlaced video to film, the decision switching is made after a delay of a period typically spanning 5 to 64 fields. This means that the encoding of the film frames in this delay period is still done in interlace mode and redundant fields in this period are not removed before encoding. Similarly when there is a transition from film to interlaced video, the interlaced video frames in the decision switching delay period are still encoded as progressive frames.

A film sequence is often being edited, and a scene change may occur in any field. Sub-titles might also be added to any field of the film, thereby changing the 3:2 repeat-field pattern of the film so that the frames are not always progressive. Interlaced video sequences also consist of some progressive frames due to very little or no motion in between these fields. The current film detection methods which have no automatic interlace/progressive detection will not be able to detect these interlaced frames within a film and the progressive frames within the interlace video.

It is therefore an object of the present invention to address the above-mentioned problems by detecting whether a frame is interlace or progressive immediately after receiving the frame data so that the encoder can encode the frame as interlace or progressive according to the detection decision, or to at least provide a useful alternative.

For existing automatic interlace and progressive detection methods, which compare the intra-field and inter-field differences to make the detection decision, the comparison is not always accurate. The inaccuracy can be due to the inter-field difference being very small, because of little or no motion between successive frames, or to the intra-field difference being large because of very detailed texture or information within the field.

There are also inaccuracy problems in detection methods which assume that interlace

difference is significantly greater than progressive difference. The problem which arises from this assumption is that when the previous field (f_{N-1}) and current field (f_N,) have little or no motion, the interlaced field difference between f_{N-1} and f_N might not be significantly greater than the difference between the progressive fields f_N and f_{N+1}.

The present invention is also intended to improve the accuracy of the interlace/progressive detection by making the detection decision which is not only based on the comparison between the interlace difference and the progressive difference, but also on the moving activities between successive frames. This is to check if an insignificant field difference between f_{N-1} and f_N is due to little motion, so as to avoid an incorrect decision due to the insignificant interlace difference.

The present invention provides a method of processing video data to detect field characteristics of the data, said data having a plurality of fields, including the steps of:

comparing first and second fields, said first field being a successive field of said second field;

comparing pixel values of respective sub-blocks of said first field and a third field, said second field being a successive field of said third field;

determining whether said first field is an interlaced field or a progressive field with respect to a successive field of said first field based on said steps of comparing.

The present invention further provides an apparatus for processing video data to detect field characteristics of the data, said data having a plurality of fields, including:

first comparison means for comparing first and second fields, said first field being a successive field of said second field;

second comparison means for comparing pixel values of respective sub-blocks of said first field and a third field, said second field being a successive field of said third field;

progressive/interlace decision means for determining whether said first field is an interlaced field or a progressive field with respect to a successive field of said first field based on respective outputs of said first and second comparison means.

A preferred embodiment of the present invention is described hereinafter, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a block diagram of a system for determining interlace/non-interlace frames, identifying repeat fields and detecting scene-changes from a video source in accordance with an embodiment of the present invention;

Figure 2 is a flow diagram illustrating the field grouping decision process;

Figure 3 is a block diagram of the consecutive field difference operation;

Figure 4 is a spatio-temporal pixel diagram illustrating the consecutive field difference computation;

Figure 5 is a flow diagram of the Interlace/progressive decision making algorithm;

Figure 6 is a spatio-temporal pixel diagram illustrating the moving region detection method.

In the preferred embodiment of the present invention, only two field memory units 101 and 102 are required. Referring to Figure 1, at a particular time, a video source 100 provides a field N to field memory 101, subtracter 103 and the consecutive field difference unit 106. At that time, the field memory 101 outputs the previous field N-1 to the second field memory 102 and to the consecutive field difference unit 106. Also at that time, the second field memory 102 outputs field N-2 to the subtracter 103. The sub-block sum of absolute differences between the pixels of the incoming fields N and N-2 (functionally expressed as *SBD (N-2,N)*), is measured using subtracter 103 and sub-block accumulator 104. The consecutive field difference between the current field N and the previous field N-1,(functionally expressed as *CFD(N-1,N)*), is measured by the consecutive field difference unit 106 and fed into an interlace/progressive decision unit 107. The value of *SBD(N-2,N)* is used in a scene change decision unit 108 to decide if field N is a new scene compared to field N-2. It is also used in a repeat field decision unit 105 to decide if field N is a repeat field of field N-2. The number of sub-block moving pixels between field N-2 and N (functionally expressed as *moving-pixel (N-2,N)*), is computed by a sub-block moving pixel counter 109, and is input to the repeat filed decision unit 105 and the interlace/progressive decision unit 107. When field N+1 arrives, *CFD(N,N+1)* is then measured and compared with *CFD(N-*

1,N) in the interlace/progressive decision unit 107. The number of sub-block moving pixels, *moving-pixel(N-1,N+1)*, is used in the interlace/progressive decision unit 107 to decide if fields N and N+1 are interlaced or progressive. The field grouping decision is made in a field grouping decision unit 110. The flow diagram relating to the field grouping decision unit 110 is shown in Figure 2. Fields N and N+1 are grouped as an interlaced or progressive frame depending on the output of the decision unit 107. If the current field N and field N+1 are detected as being interlaced by the unit 107, then fleds N and N+1 are grouped as an interlaced frame and field N+2 becomes the new current field. If fields N and N+1 are detected as being progressive, and fields N and N+2 are not detected as being repeated by unit 105, then fields N and N+1 are grouped as progressive and field N+2 becomes the new current field. However, if fields N and N+1 are detected as being progressive, and fields N and N+2 are detected as being repeated, then fields N and N+1 are gouped as being progressive, field N+2 is discarded and flied N+3 is set as the new current field.

Preferably, for all the sub-block measurements, each field is divided into 32 equal sub-blocks.

The block diagram of the consecutive field difference unit 106 is illustrated in figure 3. Subtractors 300 and 303 are used to compute the absolute pixel differences between fields N and N-1, and the smaller of the pixel differences is chosen by a comparator 301. The smaller pixel difference is then set to zero by noise attenuator 302 if it is less than a threshold T_{noise}, and each unattenuated pixel difference is accumulated in accumulator 305. This is illustrated in figure 4, where A is a pixel of the current field N and B and C are pixels from the previous field N-1 with vertical positions as shown. The pixel difference (PD) of pixel A is defined as the lesser of the absolute difference between A and B and the absolute difference between A and C, ie.

$$PD = Min(|A - B|, |A - C|)$$

The PD of every pixel in field N is computed and the values of PD less than T_{noise} are regarded as noise and set to zero. The consecutive field difference *CFD(N-1,N)*, of field f_{N-1} and field f_N, is defined as the sum of all the PDs in field f_N. The reason for selecting the

lesser of the two differences is that this will reduce inaccuracies in the calculation of the field differences arising from abnormal vertical displacement or horizontal edges. To decide whether field f_N and f_{N+1} are interlaced or progressive, the computation of the *CFD(N-1,N)* and *CFD(N,N+1)* is required.

The number of sub-block *'moving pixels'* between fields f_{N-1} and f_{N+1} is also computed by the sub-block moving pixel counter 109 to find out if there is significant motion between fields f_{N-1} and f_{N+1}. The *moving-pixel(N-1,N+1)* is defined as the pixel in each sub-block (preferably 32 sub-blocks per field) between field f_{N-1} and f_{N+1} with pixel-to-pixel difference greater than a threshold T_{move}.

A decision-making flow diagram is shown in figure 5. A ratio of *CFD(N-1,N)* to *CFD(N,N+1)* smaller than threshold T1 at step 403 indicates that fields f_N and f_{N+1} are interlaced, but to make sure that a small value of *CFD(N-1,N)* is not due to little or no motion, it is also required that the number of moving pixels between field f_{N-1} and f_{N+1} is more than threshold T2 in step 405. The decision at steps 407 and 408, as to whether the fields N and N+1 are progressive, also depends on the CFD computed for the previous frame during the decision. Prev_CFD(N,N+1) is the 'CFD(N,N+1)' computed for the previous frame (equivalent to either CFD(N-2,N-1) if the field f_{N-1} is not a repeated field or CFD(N-3,N-2) if the field f_{N-1} is detected as a repeated field). The two thresholds T3 and T4 are used to set the sensitivity of decision switching from progressive-to-interlace and interlace-to-progressive respectively (at steps 409 and 410). This is to avoid the problem of an interlaced sequence which has little or no motion switching the decision too frequently between interlace and progressive. Suitable values for T_{noise}, T_{move}, T1, T2, T3 and T4 have been found to be around 5, 30, 1.4, 100, 1.1 and 1.7 respectively.

If there is a scene change between f_{N-2} and f_N at step 401, then it may be meaningless to compare *CFD(N-1,N)* to *CFD(N,N+1)* as the scene change may occur between f_{N-1} and f_N, causing the value of *CFD(N-1,N)* to be arbitrary. The decision can only be based on the information in fields f_N and f_{N+1}. Therefore when there is a scene change detected (between

current field f_N and second previous field f_{N-2}), then the moving region detection (MRD) method is used at step 402. The MRD method detects any 'jagged region' or 'moving region' which is noticeable when two 'moving' consecutive fields are interlaced and viewed as a frame.

Referring now to figure 6 (which illustrates the MRD method), A and B are pixels from field N and C and D are pixels from field N+1 with vertical positions as shown. If the absolute difference between A and C, B and C, and B and D are all greater than a threshold $T_{interlace}$ then the pair of pixels C and D are said to be '*interlaced pixels*'. To decide whether the whole frame is interlaced, the detection is again preferably based on 32 sub-blocks. For each sub-block, if more than T_{region} number of the above '*interlaced pixels*' are detected, then the block is considered to be interlaced. If more than one block is found interlaced, then the frame is considered as interlaced.

Repeat field detection is performed on a pair of fields of the same parity (odd or even). The field similarity measurement is again preferably based on 32 sub-blocks in which the absolute sum of all the pixel-to-pixel differences of each block is accumulated in the accumulator 104. The repeat-field decision unit 105 operates as follows: The pixel differences for each sub-block difference (SBD) are compared to a threshold T_{repeat}, ie.,

$$SBD/(block\_width \times block\_height) < T_{repeat} \qquad \text{for all sub-blocks}$$

If the pixel differences are smaller than T_{repeat} for all 32 of the sub-blocks, then a repeat field is said to be detected and can be skipped for encoding by the field grouping decision unit 110. It should be noted that the repeated field detection is performed only when the incoming frame is detected as progressive by the interlace/progressive detection.

To prevent an incorrect consecutive repeat field being detected due to very little motion, the following algorithm is implemented:

If ((curr_decision=repeat-field) && (prev_decision1 = repeat-field) &&
(prev_decision3=repeat-field) && (scene-change=No))
then curr_decision=no-repeat-field
else *if* ((curr_decision=repeat-field) && (prev_decision1=repeat-field) &&
(moving-pixel > 35 in any one of the 32 sub-blocks))
then curr_decision=no-repeat-field

where *prev_decision1* is the first previous decision for repeat field detection and *prev_decision3* is the third previous decision; *scene-change* is the scene change detection decision; and *moving-pixel* is the number of pixels with pixel difference greater than T_{move} computed in the sub-block moving pixel counter 109. A suitable value for T_{repeat} has been found to be around 2.5.

The differences between the current field and the previous field of the same parity are used to detect any significant change of scene. Making use of the sub-block difference (SBD), a simple thresholding method is employed by the scene change decision unit 108. Each block difference per pixel is compared with a threshold T_{scene}. If more than T_{block} of the sub-blocks has its difference per pixel greater than T_{scene}, then a scene change is detected, ie.

$$SBD/(block\_width \times block\_height) > T_{scene} \qquad \text{for more than } T_{block} \text{ sub-blocks}$$

Apart from the above detection, a scene change is also detected by comparing the current field difference with the previous field difference to see if the current field difference has a sudden increment due to a scene change. The field difference (FD) is the sum of all the 32 absolute block differences. If the current field difference is more than T_{ratio1} times greater than the previous field difference (prev_FD), then a scene change is said to be detected. The pseudocode of the scene change detection algorithm is as follows:

While (not end of sequence)
{

```
For i = 0 to 31,
        {
        If (SBD_i/(block_widthxblock_height) > T_scene) then count = count + 1
        FD = FD + SBD_i
        }
If (count > T_block) then
        scene-change is detected
Else If (FD/prev_FD > T_ratio1) then
        scene-change is detected
If ((prev_scene_change = Yes) || (prev_FD/FD < T_ratio2)) then
        prev_FD = FD
FD = 0
Increment to next frame
}
```

In a 3:2 pulldown film sequence, subtitles may be added to a repeated field, resulting in the field not being detected as a repeat field. When this particular field becomes the current field, the current FD computed (between the current field N and second previous field N-2) will have a small value (because of the small change due to the subtitles). Therefore, in updating the previous field difference (*prev_FD*), the condition '$prev\_FD/FD < T_{ratio}$' is to avoid updating a 'repeat field difference' which will affect the scene change decision made later.

The *prev_scene-change* is a scene change decision of a previous frame. When there is a scene change detected in the previous frame, then the condition '$prev\_FD/FD < T_{ratio2}$' might not be true due to the large value of *prev-FD* and hence the criteria '*prev_scene_change = Yes*' will force an update of *prev_FD*. Suitable values for T_{scene}, T_{block}, T_{ratio1} and T_{ratio2} have been found to be about 15, 25, 2.5 and 3.0 respectively.

An advantage of embodiments of the present invention is to make accurate decisions as to whether a frame should be encoded as an interlace or progressive frame immediately after the

second field of the frame is received. This enables the MPEG encoder to encode the frame as interlace or progressive accordingly accurately, including those odd interlaced frames within a film sequence due to editing or the odd progressive frames within an interlaced video sequence. In the above-described interlace/progressive determination method, apart from comparing the consecutive field differences, the moving activities between two successive frames is also computed to ensure that interlaced fields with little or no motion will not cause an incorrect decision. The present invention also addresses the situation where the scene change occurs in the current frame. The moving region detection method is then used for the interlace/progressive determination.